

Jukebox: A Generative Model for Music

Year: 2020 | Citations: 888 | Authors: Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford

Abstract

We introduce Jukebox, a model that generates music with singing in the raw audio domain. We tackle the long context of raw audio using a multi-scale VQ-VAE to compress it to discrete codes, and modeling those using autoregressive Transformers. We show that the combined model at scale can generate high-fidelity and diverse songs with coherence up to multiple minutes. We can condition on artist and genre to steer the musical and vocal style, and on unaligned lyrics to make the singing more controllable. We are releasing thousands of non cherry-picked samples at this [https URL](#), along with model weights and code at this [https URL](#)